

Bound on Bell Inequalities by Fraction of Determinism and Reverse Triangle Inequality

P. Joshi^{1,2}, K. Horodecki^{1,3}, M. Horodecki^{1,2}, P. Horodecki⁴, R. Horodecki¹, Ben Li⁵, S. J. Szarek^{5,6} and T. Szarek⁷

¹*National Quantum Information Center of Gdańsk, 81-824 Sopot, Poland*

²*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, Gdańsk, Poland*

³*Institute of Informatics, University of Gdańsk, 80-952 Gdańsk, Poland*

⁴*Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-233 Gdańsk, Poland*

⁵*Department of Mathematics, Case Western Reserve University, Cleveland, Ohio 44106-7058, U.S.A.*

⁶*Institut de Mathématiques de Jussieu-PRG, Université Pierre et Marie Curie-Paris 6, 75252 Paris, France and*

⁷*Institute of Mathematics, University of Gdańsk, 80-952 Gdańsk, Poland*

(Dated: February 11, 2015)

It is an established fact that entanglement is a resource. Sharing an entangled state leads to non-local correlations and to violations of Bell inequalities. Such non-local correlations illustrate the advantage of quantum resources over classical resources. Here, we study quantitatively Bell inequalities with $2 \times n$ inputs. As found in [N. Gisin et al., Int. J. Q. Inf. 5, 525 (2007)] quantum mechanical correlations cannot reach the algebraic bound for such inequalities. In this paper, we uncover the heart of this effect which we call the *fraction of determinism*. We show that any quantum statistics with two parties and $2 \times n$ inputs exhibits nonzero fraction of determinism, and we supply a quantitative bound for it. We then apply it to provide an explicit *universal upper bound* for Bell inequalities with $2 \times n$ inputs. As our main mathematical tool we introduce and prove a *reverse triangle inequality*, stating in a quantitative way that if some states are far away from a given state, then their mixture is also. The inequality is crucial in deriving the lower bound for the fraction of determinism, but is also of interest on its own.

I. INTRODUCTION

Since Bell's paper [1] entanglement has been studied and explored in depth. Saying that quantum information branch emerged from extensive studies of phenomenon of entanglement would not be an exaggeration. Entanglement has been used in many information-processing applications in which it either yields an advantage over the classical setting, e.g., in communication complexity [2], or where a classical counterpart simply doesn't exist, e.g., in quantum key distribution (QKD) [3], its device independent variant (DIQKD) [4], teleportation, super dense coding [5], or Pseudo-Telepathy (PT) [6, 7].

Although quantum theory allows for violations of Bell inequalities (BI), in certain cases the violations can not reach their maximum algebraically possible value. Tsirelson was the first to find such upper bounds on Bell values for quantum theory [8] and to relate them to the Grothendieck's inequality. Much research has been done to explain why quantum mechanics does not lead to "algebraic" violations of Bell inequalities [9, 10]. In [11], Wehner and Oppenheim argue that the trade-off between steerability and uncertainty determines how non-local a theory is. In [12], Cleve et al. gave an upper bound for the winning probability for XOR games in the quantum setting; their bound depends on the classical winning probability and the Grothendieck's constant. Note that XOR game is a non-local game and that non-local games form a subset of general Bell inequalities [13].

The approach to bounding quantum violations via a Grothendieck-type constant is now quite common and reasonably well understood. It leads to estimates for Bell values that are of the form $\beta_{qm} \leq K_G \beta_{cl}$ [14]. In

this work we develop a different strategy, where the Bell value for a given inequality depends on the difference between the maximal algebraic value (β_{alg}^{max}) and maximal deterministic value (β_{det}^{max}) of the inequality in question. Specifically, we study quantitatively Bell inequalities with $2 \times n$ inputs (henceforth $2 \times n$ BI) and give a *universal bound* on quantum Bell values of these inequalities. To find this bound for $2 \times n$ BI, we introduce notion of *fraction of determinism* (FOD) and show that it depends only on the number of outcomes Alice and Bob have at their sites. We claim that presence of FOD prevents quantum Bell value from attaining the maximal algebraic value of a Bell type inequality. Our paper is inspired by Gisin et al. [15], which studied certain Bell inequalities (Pseudo-telepathy) for which quantum resources achieve algebraic violation. They show that to achieve such violations for these inequalities at minimum 3×3 inputs are required. In other words, there is no $2 \times n$ BI for which quantum theory attains algebraic violation. Here we uncover the heart of this effect – the fraction of determinism – and are able to give a quantitative bound for it.

While looking for a lower bound for FOD, we proved a fundamental property of quantum states which is interesting on its own. Namely, if ρ_1 and ρ_2 are far from σ , then any convex mixture of them is also far from σ . More precisely, if $\Delta_1 = \|\rho_1 - \sigma\|_1 \geq 2 - \epsilon$ and $\Delta_2 = \|\rho_2 - \sigma\|_1 \geq 2 - \epsilon$ for some $\epsilon \geq 0$, then, for all $p \in [0, 1]$,

$$\Delta = \|p\rho_1 + (1-p)\rho_2 - \sigma\|_1 \geq 2 - \mathcal{O}(\sqrt{\epsilon}). \quad (1)$$

where $\|\rho\|_1 \stackrel{\text{def}}{=} \text{Tr} \sqrt{\rho^\dagger \rho}$. This inequality is in a sense dual to triangle inequality since it bounds the trace distance

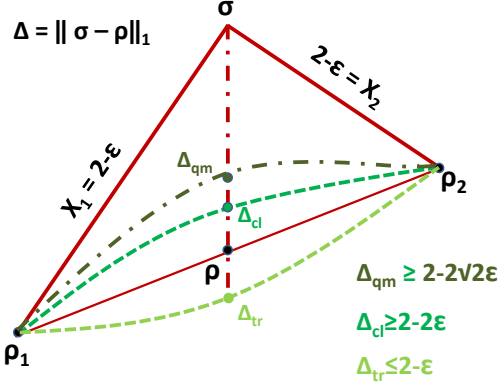


FIG. 1: Pictorial representation of different bounds of $\|\sigma - \rho\|_1$. Triangle inequality gives an upper bound $2 - \epsilon$, whereas *reverse triangle inequality* gives lower bounds $2 - 2\sqrt{2}\epsilon$ for general quantum states and $2 - 2\epsilon$ for classical (or commuting) states.

between ρ and σ from below. Accordingly, we call it “reverse triangle inequality” (RTI). Interestingly, it turns out that for classical states (commuting density matrices) one can find lower bound of Δ with the defect term linearly dependent on ϵ , while for non-commuting quantum states one can not in general have dependence better than $\mathcal{O}(\sqrt{\epsilon})$.

The second fundamental property which is used here is related to so called steering [16]. Namely, by making measurement on one site of the entangled state, one can create only those ensembles which give rise to the same density matrix – the reduced state of the entangled state. This implies that if we consider two such ensembles, there must be at least two elements (one from one ensemble, and the other from the second ensemble that are not perfectly distinguishable. It has been apparently not studied to what extent they have to be indistinguishable. Here, by using the reverse triangle inequality, we are able to give a robust quantitative bound (lemma 2), which is independent of the dimension of the underlying Hilbert space. We shall use it further to give bound for FOD, which in turn will allow to bound quantum violations for $2 \times n$ Bell inequalities.

The paper is organized in the following manner. In section II, we introduce necessary definitions and the concept of FOD. In sections III and IV, we present respectively a summary of our main results and sketches of their derivations. A special case when Bob has two inputs at his site with binary outcomes is analyzed in section V. For this special case, we have explicitly calculated a bound for FOD and for the classical fraction. Finally we conclude our work in section VI. Details of most proofs are relegated to the Appendices.

II. PRELIMINARIES

A. Definitions

Box: Consider two distant parties, Alice and Bob, sharing a physical system. Each of them perform measurements labeled as x and y respectively. Their corresponding outcomes are labeled as a and b . Then, a box is defined as family of joint probability distributions $p(a, b|x, y)$, i.e., $P = \{p(a, b|x, y)\}$. By a *non-signalling box* (NS-box) we mean a box which satisfies following conditions,

$$\begin{aligned} p(b|y) &= \sum_a p(a, b|x, y) = \sum_a p(a, b|x', y) \quad \forall b, x, x' \text{ and } y \\ p(a|x) &= \sum_b p(a, b|x, y) = \sum_b p(a, b|x, y') \quad \forall a, x, y' \text{ and } y \end{aligned} \quad (2)$$

A *local box* is defined as a box where joint probabilities can be expressed as

$$p(a, b|x, y) = \int_{\Lambda} q(\lambda) p(a|x, \lambda) p(b|y, \lambda) d\lambda, \quad (3)$$

where the hidden variable λ is distributed according to some probability density $q(\lambda)$. Such boxes satisfy (by definition, see below) every Bell inequality. We say that a box is a *Quantum box* (QM box) when conditional probabilities are realized by $p(a, b|x, y) = \text{tr}(M_a^x \otimes N_b^y \rho_{AB})$, where ρ_{AB} is a shared quantum state between party A and B, and M_a^x and N_b^y are measurements for A and B respectively. Note that, for each input x and y , $\{M_a^x\}, \{N_b^y\}$ are POVMs. In this work, we consider only NS-boxes.

Bell Inequalities: Let $S \equiv \{s_{x,y}^{a,b}\}$ be a real vector and $P = \{p(a, b|x, y)\}$ be a box then, $S \cdot P \leq \beta_T$ is called *Bell inequality* when this inequality is satisfied by any local box P [13]. Note that we can rescale the inequality and make S positive real.

Fraction of determinism (FOD): Consider a non-signalling box P . One can always express it as convex combination of $P = (1-c)X + cD$, where X is an NS-box and D is a deterministic box. The *fraction of determinism* of P is defined as

$$FOD := \max_{D, X} \{c \mid P = (1-c)X + cD\} \quad (4)$$

Classical Fraction (CF): A non-signalling box P , can always be expressed as a convex combination of $P = (1 - \sum_i c_i)X + \sum_i c_i D_i$, where X is an NS-box and D_i are deterministic boxes. Let $c_{cf} = \sum_i c_i$ then the classical fraction of a box P can be obtained by taking maximum of c_{cf} over all possible decompositions of the above form, i.e.,

$$CF(P) = \max_{\{D_i\}, X} \{c_{cf} \mid P = (1 - \sum_i c_i)X + \sum_i c_i D_i\} \quad (5)$$

Note that FOD, CF and the cost of non locality c_{nl} [17] satisfy the following relations

$$FOD \leq CF = 1 - c_{nl} \quad (6)$$

B. The Role of the Fraction of Determinism

In the classical theory, FOD is not 1. Indeed, consider the maximally mixed state, which has the smallest fraction of determinism: it is $1/k_A k_B$ where k_A is number of Alice's outcomes, and k_B is number of Bob's outcomes (assuming that the number of outcomes is same for all observables). In the quantum case, the set of states is larger, hence we might in principle have states with zero fraction of determinism. However this is not the case as shown here. On the other hand, PR-boxes [9] are completely noiseless and they do not have any fraction of determinism. The latter is equivalent to saying that they provide perfectly secure correlations. Indeed, the fraction of determinism is at the same time the fraction that can be known by the third party, or equivalently one can say that FOD in a given theory restricts Bell value from reaching its maximal algebraic value. We present the following proposition which indeed captures this idea.

If a box has some fraction c of "determinism", then this fraction implies a bound on the maximal value of a linear function (in particular for Bell type inequalities).

Proposition 1 Consider a box $P = \{p(a, b|x, y)\}$ with inputs $x \in \{x_1, \dots, x_n\}$ on Alice side and $y \in \{y_1, \dots, y_m\}$ on Bob's side. Suppose that we can find indices $a_0^{(1)}, \dots, a_0^{(n)}, b_0^{(1)}, \dots, b_0^{(m)}$ such that

$$\forall i, j \quad p(a_0^{(i)}, b_0^{(j)}|x_i, y_j) \geq c \quad (7)$$

Then for any linear function β of the box, we have

$$\beta(P) \leq \beta_{alg}^{\max} - c(\beta_{alg}^{\max} - \beta_{det}^{\max}) \quad (8)$$

where β_{alg}^{\max} is the maximum value over all boxes, while β_{det}^{\max} is the maximum over all classical deterministic boxes.

This follows from the fact that any such box can be expressed as convex combination of a deterministic box and some other box, i.e., $P = cD + (1-c)X$ and simply taking the maximal value of β .

III. SUMMARY OF THE RESULTS

We give a *universal bound* on $2 \times n$ Bell inequalities. Specifically, our main result is finding a bound on FOD for the $2 \times n$ BI scenario and showing that it only depends on number of outcomes of both parties. From Proposition 1 we know that this gives a universal bound for any linear function. A summary of our main results is as follows.

Theorem 1 For input $2 \times n$, the fraction of determinism for QM box is bounded by the following quantity:

$$c \geq \frac{0.1134}{2k l l_1 l_2} \quad (9)$$

Here, $k = \max\{|x_1|, \dots, |x_n|\}$ and $l = \max\{l_1 = |y|, l_2 = |y'|\}$, where $\{x_1, \dots, x_n\}$ are inputs on Alice's side while $\{y, y'\}$ are inputs on Bob's side.

Here, by $|z|$ we denote number of outcomes an observable z takes.

To prove the above theorem we need the following fundamental property of quantum states.

Theorem 2 (Reverse Triangle Inequality) Let $\epsilon \geq 0$ and assume that the states ρ_i, σ satisfy

$$\|\rho_i - \sigma\| \geq 2 - \epsilon \quad (10)$$

for $i = 1, \dots, l$. Then, for any probability distribution $\{p_i\}_{i=1}^l$,

1) For any states ρ_i, σ satisfying (10)

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - 2\sqrt{l}\epsilon \quad (11)$$

2) For commuting states ρ_i, σ satisfying (10)

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - l\epsilon \quad (12)$$

3) There exist three non-commuting states ρ_1, ρ_2 and σ satisfying (10) such that

$$\left\| \frac{\rho_1 + \rho_2}{2} - \sigma \right\| \leq 2 - \sqrt{2}\epsilon \quad (13)$$

Remark: The third assertion says that, in the non-commuting case, $2 - \sqrt{2}\epsilon$ is the best possible bound one can hope to achieve. Hence, one cannot have better lower bound than $2 - \mathcal{O}(\sqrt{\epsilon})$.

Using the above results, one can find lower bound for FOD in CHSH [13] case ($k = l = 2, \beta_{alg}^{\max} = 4$ and $\beta_{det}^{\max} = 2$) as $c \geq 3.5438 \times 10^{-3}$. This results in bounding CHSH value for quantum theory.

$$\begin{aligned} \beta_{CHSH}^{qm} &\leq \beta_{alg}^{\max} - c(\beta_{alg}^{\max} - \beta_{det}^{\max}) \\ &\leq 4 - 3.5438 \times 10^{-3}(4 - 2) = 3.9929 \end{aligned} \quad (14)$$

A more direct approach gives an improved bound on FOD.

$$\beta_{FOD}(P) \leq 4 - \frac{0.1096 * 2}{4} = 3.9452 \quad (15)$$

$$\beta_{CF}(P) \leq 4 - \frac{0.1123 * 2}{4} = 3.9439 \quad (16)$$

This has been elaborated in section V.

It is interesting to note that we can also roughly estimate β of (8) to upper-bound β in the classical theory. We will get a rough estimation for CHSH (in case of the maximally mixed state $c = \frac{1}{k_A k_B} = \frac{1}{4}$)

$$\beta_{CHSH}^{cl} \leq 4 - \frac{1}{4}(4 - 2) = 3\frac{1}{2} \quad (17)$$

We realize that these are weak bounds, but the importance of this study lies in their generality: they are valid for *any* Bell inequality. In the following section we shall find a bound for c for quantum states for quantum theory and derive our main results. Most of the proofs are relegated to Appendix B. We assume that Bob has 2 observables $\{y, y'\}$, i.e., $m = 2$.

IV. FRACTION OF DETERMINISM IN QM

We start with a proposition in which we redefine FOD more explicitly for QM boxes, which will lead to a lower bound that can be used in Proposition 1.

Proposition 2 *For a QM-box with $2 \times n$ input, the following quantity c_0 satisfies (7)*

$$c_0 = \inf_{\xi, \xi', \{X_r\}} \max_{r, i, j} \min\{p_i \text{Tr}(X_r \rho_i), q_j \text{Tr}(X_r \sigma_j)\} \quad (18)$$

where the infimum is taken over all ensembles $\xi = \{(p_i, \rho_i)\}_{i=1}^{|y|}$, $\xi' = \{(q_j, \sigma_j)\}_{j=1}^{|y'|}$ satisfying

$$\sum_i p_i \rho_i = \sum_j q_j \sigma_j \quad (19)$$

and over all POVMs $\{X_r\}_{r=1}^k$, i.e., sets of operators satisfying $\sum_r X_r = I$, $X_r \geq 0$, with $k = \max\{|x_1|, \dots, |x_n|\}$.

Proof: By hypothesis, our quantum box is realized via POVMs $\{M_a^x\}$ (with $x \in \{x_1, \dots, x_n\}$) on Alice's side, two POVMs $\{N_b^y, N_{b'}^{y'}\}$ on Bob's side, and a shared quantum state ρ_{AB} . Depending on Bob's measurement choice (y or y'), an ensemble $\{p(b|y), \rho_b\}_{b=1}^{|y|}$ or $\{p(b'|y'), \sigma_{b'}\}_{b'=1}^{|y'|}$ is created at Alice's site, where $p(b|y)$ and $p(b'|y')$ are marginal conditional probabilities. Even more specifically, $p(b|y)\rho_b = \text{Tr}_B((I \otimes N_b^y)\rho_{AB})$ and similarly $p(b'|y')\sigma_{b'} = \text{Tr}_B((I \otimes N_{b'}^{y'})\rho_{AB})$. These ensembles satisfy

$$\text{Tr}_B(\rho_{AB}) = \sum_b p(b|y)\rho_b = \sum_{b'} p(b'|y')\sigma_{b'}, \quad (20)$$

i.e., a condition of the type (19). If now $\{X_r\}$ is any of Alice's POVMs (say, $\{M_a^x\}$), it is apparent that the expressions $p_i \text{Tr}(X_r \rho_i)$, $q_j \text{Tr}(X_r \sigma_j)$ coincide with the conditional probabilities $p(a, b|x, y)$, $p(a, b'|x, y')$ appearing in (7). Now pick a triplet (r, i, j) such that the probabilities of the corresponding outcomes are maximal and

one can see that these indices lead to the choices of a, b that yield (7) with $c = c_0$. ■

Next, we will give an estimate on this quantity. In this way we shall obtain a universal quantum bound for any $2 \times n$ input inequalities, in terms of difference between the classical bound and the maximal algebraic bound (8). In general, c might be zero. But we show in $2 \times n$ input boxes that indeed it is bounded away from zero. To show this, one needs to prove for some choice of i, j and for any POVM X_r , that $\text{Tr}(X_r \rho_i)$ and $\text{Tr}(X_r \sigma_j)$ are bounded away from zero. Note that this indeed happens when the POVM cannot distinguish the two states ρ_i and σ_j too well. We elaborate this point through the following lemma.

Lemma 1 *Suppose that $\|\rho - \sigma\| \leq 2 - 2k\epsilon$. Then for any POVM $\{X_r\}_{r=1}^k$ there exists an outcome r_0 such that*

$$\text{Tr}(X_{r_0} \rho) \geq \epsilon \quad \text{and} \quad \text{Tr}(X_{r_0} \sigma) \geq \epsilon \quad (21)$$

Note that using this lemma we can replace conditional probabilities by ϵ and get rid of choosing maximum for all r and the optimization over $\{X_r\}$. The above lemma asserts that there exist at least one outcome r for each input and each pair (ρ_i, σ_j) such that the corresponding probabilities are lower-bounded by ϵ_{ij} , i.e., $\text{Tr}(X_r \rho_i) \geq \epsilon_{ij}$ and $\text{Tr}(X_r \sigma_j) \geq \epsilon_{ij}$. Therefore one can simplify the expression for FOD as follows.

$$\begin{aligned} c_0 &\geq c_1 = \inf_{\xi, \xi', \{X_r\}} \max_{r, i, j} \min\{p_i \epsilon_{ij}, q_j \epsilon_{ij}\} \\ c_1 &\geq \frac{1}{2k} \inf_{\xi, \xi'} \max_{i, j} \min\{p_i (2 - \|\rho_i - \sigma_j\|), q_j (2 - \|\rho_i - \sigma_j\|)\} \end{aligned} \quad (22)$$

where we assume $\|\rho_i - \sigma_j\| \leq 2 - 2k\epsilon_{ij}$.

Having simplified FOD, we will now state and apply a theorem which is both vital for our results, as well as important on its own.

Theorem 2 *{Restatement} Let $\epsilon \geq 0$ and assume that the states ρ_i, σ satisfy*

$$\|\rho_i - \sigma\| \geq 2 - \epsilon \quad (23)$$

for $i = 1, \dots, l$. Then, for any probability distribution $\{p_i\}_{i=1}^l$,

1) *For any states ρ_i, σ satisfying (23)*

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - 2\sqrt{l}\epsilon \quad (24)$$

2) *For commuting states ρ_i, σ satisfying (23)*

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - l\epsilon \quad (25)$$

3) There exist three non-commuting states ρ_1, ρ_2 and σ satisfying (23) such that

$$\left\| \frac{\rho_1 + \rho_2}{2} - \sigma \right\| \leq 2 - \sqrt{2\epsilon} \quad (26)$$

We relegate the proof of the above Theorem to Appendix A.

Using this theorem we argue that for two ensembles (19), which give rise to the same density matrix, $\|\sigma_{i_0} - \rho_{j_0}\|$ must be bounded away from 2 for some i_0, j_0 . In general, we have the following lemma.

Lemma 2 For two ensembles $\{(p_i, \rho_i)\}_{i=1}^{|y|}, \{(q_j, \sigma_j)\}_{j=1}^{|y'|}$ satisfying

$$\left\| \sum_i p_i \rho_i - \sum_j q_j \sigma_j \right\| \leq x \quad (27)$$

there exist i_0 and j_0 such that

$$\|\rho_{i_0} - \sigma_{j_0}\| \leq 2 - \epsilon \quad (28)$$

where ϵ is solution of the following equation

$$2 - 2\sqrt{l_1 l_2 \epsilon} = x \quad (29)$$

where $|y| = l_1$ and $|y'| = l_2$.

We are now almost done. However, it may still happen that, for the chosen pair of indices, the probabilities p_{i_0}, q_{j_0} are small, and we will not have a bound for the whole quantity of (22). Therefore we need to truncate the ensembles so that the minimal probability is bounded away from zero. Such smaller ensembles, do not give rise anymore to the same density matrix. However their density matrices are still close, provided we have not truncated too much.

Lemma 3 Suppose we are given two ensembles

$$\mathcal{E}_1 = \{p_i, \rho_i\}_{i=1}^{l_1}, \quad \mathcal{E}_2 = \{q_j, \sigma_j\}_{j=1}^{l_2} \quad (30)$$

which give rise to the same density matrix. Let p_i and q_j be arranged in the nonincreasing order. Let us denote

$$\delta_1 = 1 - \sum_{i=1}^{\tilde{l}_1} p_i, \quad \delta_2 = 1 - \sum_{j=1}^{\tilde{l}_2} q_j \quad (31)$$

Consider new ensembles

$$\tilde{\mathcal{E}}_1 = \{\tilde{p}_i, \rho_i\}_{i=1}^{\tilde{l}_1}, \quad \tilde{\mathcal{E}}_2 = \{\tilde{q}_j, \sigma_j\}_{j=1}^{\tilde{l}_2} \quad (32)$$

where $\tilde{p}_i = p_i/(1 - \delta_1), \tilde{q}_j = q_j/(1 - \delta_2)$. Then the new ensembles satisfy

$$\left\| \sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j \right\| \leq \frac{2 \max\{\delta_1, \delta_2\}}{1 - \min\{\delta_1, \delta_2\}} \quad (33)$$

Thus we can use the new ensembles to show that there exist a pair of states ρ_{i_0} and σ_{j_0} , and that at the same time the weights of the states satisfy $p_{i_0} \geq p_{\tilde{l}_1}, q_{j_0} \geq q_{\tilde{l}_2}$. Thus adjusting \tilde{l}_1 and \tilde{l}_2 properly we can simultaneously secure a bound on both the weights and the norm.

We can now prove our final result.

Theorem 1 {Restatement} For input $2 \times n$, the fraction of determinism for QM box is bounded by the following quantity

$$c \geq \frac{0.1134}{2k l l_1 l_2}$$

Here, $k = \max\{|x_1|, \dots, |x_n|\}$ and $l = \max\{|y|, l_2 = |y'|\}$, where $\{x_1, \dots, x_n\}$ are inputs on Alice's side while $\{y, y'\}$ are inputs on Bob's side.

Proof. First truncate the ensembles appropriately. We use the notation of Lemma 3. Let $\mu > 1$ be a parameter. Let us choose the largest \tilde{l}_1 and \tilde{l}_2 such that $p_{\tilde{l}_1} > \frac{1}{l\mu}$ and $q_{\tilde{l}_2} > \frac{1}{l\mu}$, where $l = \max\{l_1, l_2\}$. Then δ_1 and δ_2 not larger than $l \times \frac{1}{l\mu}$. Consequently, we get the following estimate on the truncated ensemble:

$$\left\| \sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j \right\| \leq \frac{2/\mu}{1 - 1/\mu} = \frac{2}{\mu - 1}. \quad (34)$$

From equation (22) and from Lemma 2 it follows that

$$c \geq \frac{1}{2k} \frac{\epsilon}{l\mu}, \quad (35)$$

where ϵ satisfies equation (29) with $x = \frac{2}{\mu-1}$. After some simplifications, we get

$$\epsilon \geq \left(\frac{(\mu - 2)}{l(\mu - 1)} \right)^2. \quad (36)$$

Substituting this ϵ value in the preceding equation we are led to

$$c \geq \frac{1}{2kl^3} \left(\frac{(\mu - 2)}{(\mu - 1)} \right)^2 \frac{1}{\mu}. \quad (37)$$

We now note that the function $f(\mu) = \frac{1}{\mu} \left(\frac{\mu-2}{\mu-1} \right)^2$ reaches its maximum at $\mu_0 = (5 + \sqrt{17})/2$, which completes the proof. ■

Example: Consider the CHSH case, where $k = 2, l = 2$ and substituting these values we find $FOD \geq 7.0875 \times 10^{-3}$. Consequently

$$\beta_{CHSH}^{qm} \leq \beta_{alg} - c(\beta_{alg} - \beta_{det}) = 3.9858. \quad (38)$$

In the next section bounds for FOD and CF are calculated for a simple case of $2 \times n$ input with binary outcomes on Bob's side. One can find these bounds using some of the lemmas and propositions described in section III, which in turn gives an even better bound than the ones obtained using the general result of Theorem 1.

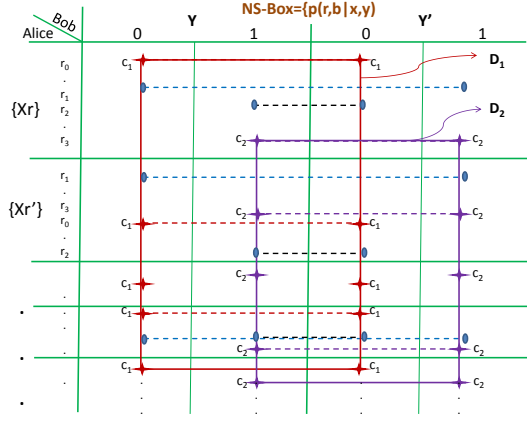


FIG. 2: The box $P = \{p(r, b | X_r, y)\}$ of Alice and Bob. D_1 and D_2 are two orthogonal deterministic boxes with fraction c_1 and c_2 respectively. And these can be subtracted from P .

V. FOD AND CF FOR BINARY OUTCOMES ON BOB'S SIDE

Using structural property of boxes and Lemmas 1 and 3 and Theorem 2 in section IV, one can explicitly compute bounds for FOD and CF for the case when Bob has binary outcomes. Technically, we look for structures of deterministic boxes within the structure of the quantum box. The maximum fraction of these deterministic boxes bound FOD of the quantum box. This technique is explained below and on Fig. 2.

Bob can create $\{p_i \rho_i\}_{i=0}^1$ or $\{q_j \sigma_j\}_{j=0}^1$ ensemble at Alice's site by making measurement y or y' respectively on his part of shared quantum state. Lemma 1 asserts that for all pairs of ρ_i and σ_j and for all POVMs $\{X_r\}$

$$\begin{aligned} \exists \quad & \epsilon_{ij} \geq 0, X_{r_0}, X_{r_1}, X_{r_2}, X_{r_3} \text{ such that} \\ & \text{tr}(X_{r_0} \rho_0) \geq \epsilon_{00}, \text{ and } \text{tr}(X_{r_0} \sigma_0) \geq \epsilon_{00} \\ & \text{tr}(X_{r_1} \rho_0) \geq \epsilon_{01}, \text{ and } \text{tr}(X_{r_1} \sigma_1) \geq \epsilon_{01} \\ & \text{tr}(X_{r_2} \rho_1) \geq \epsilon_{10}, \text{ and } \text{tr}(X_{r_2} \sigma_0) \geq \epsilon_{10} \\ & \text{tr}(X_{r_3} \rho_1) \geq \epsilon_{11}, \text{ and } \text{tr}(X_{r_3} \sigma_1) \geq \epsilon_{11}, \end{aligned} \quad (39)$$

where $\epsilon_{ij} \leq \frac{1}{2k}(2 - \|\rho_i - \sigma_j\|)$. This means that when Bob obtains outcomes (b, b') for inputs (y, y') then for any POVM of Alice there exist at least one outcome, call it a *confusing outcome*, on her side such that once she obtains it, she cannot distinguish between measurement choices of Bob with certainty, i.e., to determine whether Bob chose y or y' to create the first ensemble. For example, in the first pair of inequalities in (39) above, the outcome r_0 of some POVM can not tell apart with certainty ρ_{00} from σ_{00} . There are four pairs of (b, b') , hence there are four confusing outcomes corresponding to each of these four cases.

Consider the particular case when Bob obtains $(0, 0)$ when he measures (y, y') , and let us say r_0 is a confusing outcome for Alice when she chooses to measure

POVMs $\{X_r\}$. Since Bob obtains $(0, 0)$, the marginals satisfy $p_0 > 0$ and $q_0 > 0$. Lemma 1 asserts that for any measurement choice we have $\text{Tr}(X_{r_0} \rho_0) \geq \epsilon_{00}$ and $\text{Tr}(X_{r_0} \sigma_0) \geq \epsilon_{00}$. Hence for every POVM, there is at least one confusing outcome on Alice's side. Therefore, in the quantum box we can replace the probabilities corresponding to each of these confusing outcomes for every measurement choice of Alice with $c_{00} := \min\{p_0 \epsilon_{00}, q_0 \epsilon_{00}\}$. One can now see that by this construction we can create a deterministic box (say D_{00}) with fraction equal to c_{00} . In other words, every quantum box P_Q satisfies the relation $P_Q \geq (1 - c_{00})X + c_{00}D_{00}$. In such a way, we can create four separate deterministic boxes $(\{D_{ij}\}_{i,j=0}^1)$ corresponding to each of the outcome pairs (b, b') of Bob.

There is a possibility that there may exist a measurement setting for Alice such that she obtains a single confusing outcome for two or more different cases, e.g., when she obtains a confusing outcome r_0 , she is unable to distinguish between measurement choices of Bob not only in the case when Bob obtains $(0, 0)$ but also in the case when he obtains $(1, 0)$. So, in the worst case, for some measurement choices there may be just one confusing outcome at Alice's side for all the four different cases as in Fig. 3 in the last row of the box. In that case, the quantum box does not satisfy $P_Q \geq (1 - \sum c_{ij})X + \sum_{i,j=0}^1 c_{ij}D_{ij}$ because this would require us to use some probabilities twice. Therefore, in general, one can use only orthogonal pairs of deterministic boxes to resolve this issue, i.e., either $P_Q \geq (1 - c_{00} - c_{11})X + c_{00}D_{00} + c_{11}D_{11}$ or $P_Q \geq (1 - c_{01} - c_{10})X + c_{01}D_{01} + c_{10}D_{10}$.

The maximum fraction of such deterministic boxes bounds from below the FOD of the QM box under consideration. The sum of these fractions bounds CF. So to calculate FOD and CF for a fixed ensemble, we need to find

$$\begin{aligned} FOD = \frac{1}{2k} \max\{ & \min\{p_0 \epsilon_{00}, q_0 \epsilon_{00}\}, \min\{p_1 \epsilon_{11}, q_1 \epsilon_{11}\}, \\ & \min\{p_0 \epsilon_{01}, q_1 \epsilon_{01}\}, \min\{p_1 \epsilon_{10}, q_0 \epsilon_{10}\} \} \end{aligned} \quad (40)$$

and

$$\begin{aligned} CF = \frac{1}{2k} \max\{ & \min\{p_0 \epsilon_{00}, q_0 \epsilon_{00}\} + \min\{p_1 \epsilon_{11}, q_1 \epsilon_{11}\}, \\ & \min\{p_0 \epsilon_{01}, q_1 \epsilon_{01}\} + \min\{p_1 \epsilon_{10}, q_0 \epsilon_{10}\} \} \end{aligned} \quad (41)$$

To calculate these values w.l.g. we can assume $p_0 \leq q_0 \leq q_1 \leq p_1$. Using lemma(3) and optimizing over p 's and q 's we finally get the following values (appendix contains detailed calculations)

Using theorem 2 we find, $FOD = \frac{0.1096}{2k}$, $CF = 0.1122/2k$ and for $k=2$

$$\beta_{FOD}(P) \leq 4 - \frac{0.1096 * 2}{4} = 3.9452 \quad (42)$$

$$\beta_{CF}(P) \leq 4 - \frac{0.1123 * 2}{4} = 3.9439 \quad (43)$$

Bob Alice		y		y'	
		0	1	0	1
{Xr}	r ₀	ρ ₀ ε ₀₀		q ₀ ε ₀₀	
	r ₁	ρ ₀ ε ₀₁		q ₁ ε ₀₁	
	r ₂		ρ ₁ ε ₁₀	q ₀ ε ₁₀	
	r ₃		ρ ₁ ε ₁₁	q ₁ ε ₁₁	
{Xr'}	r ₁	ρ ₀ ε ₀₁		q ₁ ε ₀₁	
	r ₃	ρ ₀ ε ₀₀	ρ ₁ ε ₁₁	q ₀ ε ₀₀	q ₁ ε ₁₁
	r ₀		ρ ₁ ε ₁₀	q ₀ ε ₁₀	
	r ₂				
.	.	ρ ₀ ε ₀₀		q ₀ ε ₀₀	
	.	ρ ₀ ε ₀₁	ρ ₁ ε ₁₀	q ₀ ε ₁₀	q ₁ ε ₁₁
	.				
	.				

FIG. 3: The box $\{p(r, b|X_r, y)\}$ of Alice and Bob. Dashed lines represents which pairs are being confused and their lower bounds. Note that r_i 's are independent of which input Alice chooses. Some or all r_i 's may coincide with each other for some inputs of Alice.

These bounds are very weak, but since they hold for any $2 \times n$ Bell type inequalities, they presumably can not be much better than this.

VI. CONCLUSION

Here we have given quantitatively a *universal bound* for $2 \times n$ input Bell inequalities, which is independent of the number 'n' of inputs. Specifically, we show that this universal bound depends on the number of outputs of the two parties and on the difference between the maximal algebraic value and the maximal deterministic value of the

inequality. We show that presence of FOD in $2 \times n$ BI prevents quantum Bell values from achieving the maximal algebraic value. Hence this result is also a quantitative proof of the theorem shown by Gisin et al. in [15], which states that there exist no $2 \times n$ input Pseudo-Telepathy game. Although these bounds are not tight, one can improve them by considering the classical fraction and generalize the result using it. We have analyzed a simple case where the classical fraction gives better bound than taking into account merely FOD.

To obtain the above results, we established a *reverse triangle inequality*, which is an independent result of its own interest. The triangle inequality gives upper bounds on trace distance between two states, whereas RTI bounds the trace distance from below. We have determined that this bound is different for non-commuting states than when considering only commuting states. The bound in the commuting case is sharp, and the one in the non-commuting case is close to being sharp.

Acknowledgments

We thank Aram Harrow for a suggestion that lead to a simpler proof of our geometric result and to a better constant. PJ thanks P. Mazurek for useful discussions. This work was supported by ERC QO-LAPS, EC IP QESSENCE, EC grant RAQUEL, MNiSW grant IdP2011 000361 and NCBiR-CHIST-ERA Project QUASAR. PJ was also supported by grant MPD/2009-3/4 from Foundation for Polish Science. SJS was partially supported by grants from the National Science Foundation (U.S.A.) and by the grant 2011-BS01-008-02 from ANR (France). TS was partially supported by the National Science Centre of Poland, grant number DEC-2012/07/B/ST1/03320. Part of this work was done at the National Quantum Information Centre of Gdańsk.

-
- [1] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
 - [2] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Math. Phys. **82**, 665 (2010).
 - [3] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002), arXiv:quant-ph/0101098.
 - [4] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007), arXiv.org:quant-ph/0702152.
 - [5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
 - [6] G. Brassard, R. Cleve, and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).
 - [7] G. Brassard, A. Broadbent, and A. Tapp, Fortschr. Phys **35**, 1877 (2005), arXiv:quant-ph/0407221.
 - [8] B. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).
 - [9] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
 - [10] M. Horodecki, P. Horodecki, R. Ramanathan, J. Tuziemski (2014), arXiv:1410.0947.
 - [11] J. Oppenheim and S. Wehner, Science **330**, 1072 (2010), arXiv:1004.2507.
 - [12] R. Cleve, P. Hoyer, B. Toner, and J. Watrous (2004), quant-ph/0404076.
 - [13] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
 - [14] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf, Comm. Math. Phys. **300**, 715 (2010).
 - [15] N. Gisin, A. A. Methot, and V. Scarani, International Journal of Quantum Information **5**, 525 (2007), arXiv:quant-ph/0610175.
 - [16] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
 - [17] N. Brunner, D. Cavalcanti, A. Salles, and P. Skrzypczyk, Phys. Rev. Lett. **106**, 020402 (2011), arXiv:1009.4207.
 - [18] J.-C. Bourin, Proceedings of the American Mathematical Society **138**(2), 495 (2010).
 - [19] C. A. Fuchs and J. van de Graaf, IEEE Trans. Inf. Theory **45**, 1216 (1999), quant-ph/9712042.

Appendix A: Proof of Theorem 2 and a discussion of its optimality

Theorem 2 *{Restatement}* Let $\epsilon \geq 0$ and assume that the states ρ_i, σ satisfy

$$\|\rho_i - \sigma\| \geq 2 - \epsilon \quad (\text{A1})$$

for $i = 1, \dots, l$. Then, for any probability distribution $\{p_i\}_{i=1}^l$,

1) For any states ρ_i, σ satisfying (A1)

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - 2\sqrt{l\epsilon} \quad (\text{A2})$$

2) For commuting states ρ_i, σ satisfying (A1)

$$\left\| \sum_{i=1}^l p_i \rho_i - \sigma \right\| \geq 2 - l\epsilon \quad (\text{A3})$$

3) There exist three non-commuting states ρ_1, ρ_2 and σ satisfying (A1) such that

$$\left\| \frac{\rho_1 + \rho_2}{2} - \sigma \right\| \leq 2 - \sqrt{2\epsilon} \quad (\text{A4})$$

Proof. We start by recalling two well-known facts.

Rotfel'd Inequality [18]: Let f be a concave function on $[0, \infty)$ such that $f(0) \geq 0$ and let $A_1, \dots, A_l \geq 0$. Then

$$\text{Tr} f\left(\sum_{i=1}^l A_i\right) \leq \text{Tr} \sum_{i=1}^l f(A_i) \quad (\text{A5})$$

Rotfel'd Inequality is usually stated for just two matrices (i.e., $l = 2$), but the general case follows easily by induction.

Fuchs–van de Graaf inequalities [19]: These inequalities give two-sided bounds for the trace distance between two quantum states σ and τ in terms of *fidelity* between σ and τ , which is defined as $F(\sigma, \tau) = \text{Tr} \sqrt{\sqrt{\sigma} \tau \sqrt{\sigma}}$. We have

$$1 - F(\sigma, \tau) \leq \frac{1}{2} \|\sigma - \tau\| \leq \sqrt{1 - F(\sigma, \tau)^2}. \quad (\text{A6})$$

Rotfel'd inequality applied with $f(t) = \sqrt{t}$ allows us to upper-bound fidelity of the mixture $\sum_{i=1}^l p_i \rho_i =: \rho$ in terms of individual fidelities:

$$F(\sigma, \rho) = \text{Tr} \sqrt{\sqrt{\sigma} \left(\sum_{i=1}^l p_i \rho_i \right) \sqrt{\sigma}} \quad (\text{A7})$$

$$= \text{Tr} \sqrt{\sum_{i=1}^l p_i \sqrt{\sigma} \rho_i \sqrt{\sigma}} \quad (\text{A8})$$

$$\leq \sum_{i=1}^l \sqrt{p_i} \text{Tr} \sqrt{\sqrt{\sigma} \rho_i \sqrt{\sigma}}$$

$$= \sum_{i=1}^l \sqrt{p_i} F(\sigma, \rho_i)$$

The second inequality in (A6) can be rewritten as

$$F(\sigma, \tau)^2 \leq 1 - \frac{1}{4} \|\sigma - \tau\|^2, \quad (\text{A9})$$

which combined with the hypothesis $\|\rho_i - \sigma\| \geq 2 - \epsilon$ leads to

$$F(\sigma, \rho_i) \leq \sqrt{1 - \frac{1}{4}(2 - \epsilon)^2} = \sqrt{\epsilon - \frac{\epsilon^2}{4}} \leq \sqrt{\epsilon}. \quad (\text{A10})$$

Inserting this bound into (A7) and using Cauchy-Schwarz inequality yields

$$F(\sigma, \rho) \leq \sum_{i=1}^l \sqrt{p_i} \sqrt{\epsilon} \leq \sqrt{l\epsilon}. \quad (\text{A11})$$

We are now in a position to appeal to the first of the Fuchs-van de Graaf inequalities (A6) to obtain

$$\frac{1}{2} \|\rho - \sigma\| \geq 1 - F(\sigma, \rho) \geq 1 - \sqrt{l\epsilon} \quad (\text{A12})$$

or $\|\rho - \sigma\| \geq 2 - 2\sqrt{l\epsilon}$, as needed.

The dependence of the bound in (A2) on ϵ (and presumably on l) can not be significantly improved. To put this in a perspective, let us state an analogous result for classical states, i.e., probability densities (non-negative functions with unit integral).

Let g_i, h be probability densities satisfying $\|g_i - h\|_1 = \int |g_i - h| \geq 2 - \epsilon$ for $i = 1, \dots, l$. Then, for any weights $\{p_i\}_{i=1}^l$,

$$\left\| \sum_{i=1}^l p_i g_i - h \right\|_1 \geq 2 - l\epsilon \quad (\text{A13})$$

and the inequality is sharp.

Indeed, since for $u, v \geq 0$, $|u - v| = u + v - 2 \min\{u, v\}$, the condition $\|g_i - h\|_1 = \int g_i + h - 2 \min\{g_i, h\} \geq 2 - \epsilon$ translates to $\int \min\{g_i, h\} \leq \epsilon/2$. Accordingly, if $g = \sum_{i=1}^l p_i g_i$, then

$$\min\{g, h\} \leq \sum_{i=1}^l \min\{g_i, h\} \quad (\text{A14})$$

and so $\int \min\{g, h\} \leq l\epsilon/2$, which is again equivalent to $\|g - h\|_1 \geq 2 - l\epsilon$.

While the “threshold for significance” in the bounds in (A2) and (A13) is roughly the same ($l\epsilon \ll 1$), the dependence on $l\epsilon$ as that quantity goes to 0 is different. What is interesting is that this difference between the classical and quantum settings is real and not just an artifact of the argument. What follows is an example showing that the $\mathcal{O}(\sqrt{\epsilon})$ dependence in (A2) is optimal. We will focus on the case $l = 2$.

To simplify the exposition, let us first reformulate the problem by considering a slightly more general question: *What is the optimal function $\epsilon \mapsto \phi(\epsilon)$ such that whenever*

ρ_1, ρ_2, σ are positive semi-definite matrices whose trace is at most 1 and such that $\text{Tr}\rho_i + \text{Tr}\sigma - \|\rho_i - \sigma\| \leq \epsilon$ for $i = 1, 2$, then $\text{Tr}\rho + \text{Tr}\sigma - \|\rho - \sigma\| \leq \phi(\epsilon)$ for any convex combination $\rho = p\rho_1 + (1-p)\rho_2$? The point is that the optimal function ϕ for this relaxed problem is the same as for the original problem when all the traces are required to be equal to 1 at the cost of increasing the dimension by 2. Indeed, if ρ_i, σ are as above, we may define states $\tilde{\rho}_i, \tilde{\sigma}$ by

$$\tilde{\rho}_i = \begin{bmatrix} \rho_i & 0 & 0 \\ 0 & 1 - \text{Tr}\rho_i & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{\sigma} = \begin{bmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 - \text{Tr}\sigma \end{bmatrix} \quad (\text{A15})$$

It is then easy to see that $2 - \|\tilde{\rho}_i - \tilde{\sigma}\| = \text{Tr}\rho_i + \text{Tr}\sigma - \|\rho_i - \sigma\|$, and similarly for $\tilde{\rho} = p\tilde{\rho}_1 + (1-p)\tilde{\rho}_2$.

With this reformulation, it is enough to look at 2×2 matrices and $p_1 = \frac{1}{2}$. Given $r \in [0, 1]$, consider

$$\sigma = \begin{bmatrix} 0 & 0 \\ 0 & r \end{bmatrix}, \quad \rho_i = \begin{bmatrix} 1-r & \pm\sqrt{r(1-r)} \\ \pm\sqrt{r(1-r)} & r \end{bmatrix}, \quad (\text{A16})$$

where $i = 1$ corresponds to the plus sign and $i = 2$ to the minus. One directly checks that

$$\text{Tr}\rho_i + \text{Tr}\sigma - \|\rho_i - \sigma\| = 1 + r - \sqrt{1 + 2r - 3r^2} \quad (\text{A17})$$

On the other hand, if $\rho = \frac{1}{2}(\rho_1 + \rho_2)$, then

$$\text{Tr}\rho + \text{Tr}\sigma - \|\rho - \sigma\| = 2r. \quad (\text{A18})$$

In our setting, this means that if $\epsilon := 1 + r - \sqrt{1 + 2r - 3r^2}$ (which covers all possible values $\epsilon \in [0, 2]$ as r varies over $[0, 1]$), then $\phi(\epsilon) \geq 2r$. Since $(2r)^2 \geq 2(1 + r - \sqrt{1 + 2r - 3r^2}) = 2\epsilon$ for $r \in [0, 1]$, this shows that $\phi(\epsilon) \geq \sqrt{2\epsilon}$. In other words, for $l = 2$ one can not have a lower bound in (A2) that is better than $2 - \sqrt{2\epsilon}$.

While this example does not directly address the case $l > 2$, we know that – already in the classical setting – one can not have a nontrivial bound if $l\epsilon$ is not small enough, and so the dependence of the bound in (A2) on l can not be too far from optimal.

Appendix B: FOD in QM

Proposition 2 *The following quantity c_0 satisfies (7)*

$$c_0 = \inf_{\xi, \xi', \{X_r\}} \max_{r, i, j} \min\{p_i \text{Tr}(X_r \rho_i), q_j \text{Tr}(X_r \sigma_j)\} \quad (\text{B1})$$

where the infimum is taken over ensembles $\xi = \{(p_i, \rho_i)\}_{i=1}^{|y|}$, $\xi' = \{(q_j, \sigma_j)\}_{j=1}^{|y'|}$ satisfying

$$\sum_i p_i \rho_i = \sum_j q_j \sigma_j \quad (\text{B2})$$

and over all POVMs $\{X_r\}_{r=1}^k$, i.e., sets of operators satisfying $\sum_r X_r = I$, $X_r \geq 0$, with $k = \max\{|x_1|, \dots, |x_n|\}$.

Proof. We rewrite joint probabilities in terms of conditional probabilities. Clearly, for a given ensemble and a fixed input x_n , we can find two numbers (a_0, b_0) such that,

$$\begin{aligned} p_{ab|XY}(a_0, b_0) &:= \\ \max_{a, b, b'} \min\{p(b|y)p(a^{(n)}|x_n, b), p(b'|y')p(a^{(n)}|x_n, b')\} \\ &\geq \max_{r, i, j} \min\{p_i \text{tr}(X_r \rho_i), q_j \text{tr}(X_r \sigma_j)\} \end{aligned}$$

where, $p_i = p(b|y)$ and $q_j = p(b'|y')$ and $\text{tr}(X_r \rho_i)$ and $\text{tr}(X_r \sigma_j)$ are conditional probabilities. And we consider POVMs $\{X_r\}_{r=1}^k$ with $k = \max\{|x_1|, \dots, |x_n|\}$. Taking infimum over all ensembles ξ, ξ' and inputs $\{X_r\}$, gives us RHS = c_0 . Hence the proposition. ■

Lemma 1 *Suppose that $\|\rho - \sigma\| \leq 2 - 2k\epsilon$. Then for any POVM $\{X_r\}_{r=1}^k$ there exists an outcome r_0 such that*

$$\text{Tr}(X_{r_0} \rho) \geq \epsilon, \quad \text{and} \quad \text{Tr}(X_{r_0} \sigma) \geq \epsilon \quad (\text{B3})$$

Proof. One shows that if, on the contrary, for all r we have either $\text{Tr}(\rho X_r) \leq \epsilon$ or $\text{Tr}(\rho X_r) \leq \epsilon$, then

$$p_e \leq \frac{1}{2}\epsilon k, \quad (\text{B4})$$

where k is the number of outcomes of the POVM, and p_e is probability of error in distinguishing ρ versus σ with equal apriori probabilities given by Helstrom relation

$$p_e(\rho, \sigma) = \frac{1}{2} - \frac{1}{4}\|\rho - \sigma\|. \quad (\text{B5})$$

To prove it, let us define two sets: $I_\rho = \{r : \text{Tr}(\sigma X_r) \leq \epsilon\}$ and $I_\sigma = I \setminus I_\rho$ where I is the set of all indices r . By the above assumption, for all $r \in I_\sigma$ we have $\text{Tr}(\rho X_r) \leq \epsilon$. Our decision scheme will be now: if $r \in I_\rho$ then the state is ρ , otherwise it is σ . With this decision scheme we have

$$\begin{aligned} p_e &\leq \frac{1}{2}\text{Tr}\left(\sum_{r \in I_\rho} X_r \sigma\right) + \frac{1}{2}\text{Tr}\left(\sum_{r \in I_\sigma} X_r \rho\right) \\ &\leq \frac{1}{2}|I_\rho|\epsilon + \frac{1}{2}|I_\sigma|\epsilon = \frac{1}{2}k\epsilon. \end{aligned} \quad (\text{B6})$$

Thus, if $p_e \geq \frac{1}{2}\epsilon k$ then, there must exist such an outcome r_0 that both inequalities (B3) hold. ■

Lemma 2 *For two ensembles $\{(p_i, \rho_i)\}_{i=1}^{|y|}$, $\{(q_j, \sigma_j)\}_{j=1}^{|y'|}$ satisfying*

$$\left\| \sum_i p_i \rho_i - \sum_j q_j \sigma_j \right\| \leq x \quad (\text{B7})$$

there exist i_0 and j_0 such that

$$\|\rho_{i_0} - \sigma_{j_0}\| \leq 2 - \epsilon \quad (\text{B8})$$

where ϵ is solution of the following equation

$$2 - 2\sqrt{l_1 l_2 \epsilon} = x \quad (\text{B9})$$

where $|y| = l_1$ and $|y'| = l_2$.

Proof. Let us first show it for $|y| = |y'| = 2$. Let us assume on the contrary, that for all pairs (i, j) of indices $\|\rho_i - \sigma_j\| \geq 2 - \epsilon$. Then together with theorem 2 imply that

$$\|p_1\rho_1 + p_2\rho_2 - q_1\sigma_1 - q_2\sigma_2\| \geq 2 - 4\sqrt{\epsilon} \quad (\text{B10})$$

However, since the two ensembles give rise to the same density matrix, we have

$$2 - 4\sqrt{\epsilon} \leq x \quad (\text{B11})$$

This implies, that at least one of the pairs must satisfy

$$\|\rho_{i_0} - \sigma_{j_0}\| \leq 2 - \tilde{\epsilon} \quad (\text{B12})$$

where $\tilde{\epsilon}$ is solution of (B11), if we put equality.

Let us now extend the proof to the more general case. By theorem 2 there is for all $i \in \{1, \dots, |y|\}$

$$\|\rho_i - \sum_{j=1}^{|y'|} q_j \sigma_j\| \geq 2 - 2\sqrt{|y'|\epsilon} \quad (\text{B13})$$

Applying the theorem 2 again, we obtain

$$\|\sum_{i=1}^{|y|} p_i \rho_i - \sum_{j=1}^{|y'|} q_j \sigma_j\| \geq 2 - 2\sqrt{|y'|\epsilon} \quad (\text{B14})$$

hence by assumption

$$x \geq 2 - 2\sqrt{|y'|\epsilon}, \quad (\text{B15})$$

and we obtain the thesis. ■

Lemma 3 Suppose we are given two ensembles

$$\mathcal{E}_1 = \{p_i, \rho_i\}_{i=1}^{l_1}, \quad \mathcal{E}_2 = \{q_j, \sigma_j\}_{j=1}^{l_2} \quad (\text{B16})$$

which give rise to the same density matrix. Let p_i and q_j be in decreasing order. Let us denote:

$$\delta_1 = 1 - \sum_{i=1}^{\tilde{l}_1} p_i, \quad \delta_2 = 1 - \sum_{j=1}^{\tilde{l}_2} q_j \quad (\text{B17})$$

We then consider new ensembles

$$\tilde{\mathcal{E}}_1 = \{\tilde{p}_i, \rho_i\}_{i=1}^{\tilde{l}_1}, \quad \tilde{\mathcal{E}}_2 = \{\tilde{q}_j, \sigma_j\}_{j=1}^{\tilde{l}_2} \quad (\text{B18})$$

where $\tilde{p}_i = p_i/(1 - \delta_1)$, $\tilde{q}_j = q_j/(1 - \delta_2)$. Then the new ensembles satisfy

$$\|\sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j\| \leq \frac{2 \max\{\delta_1, \delta_2\}}{1 - \min\{\delta_1, \delta_2\}} \quad (\text{B19})$$

Proof. From triangle inequality

$$\begin{aligned} 0 &= \left\| \sum_{i=1}^{l_1} p_i \rho_i - \sum_{j=1}^{l_2} q_j \sigma_j \right\| \geq \\ &\left\| \sum_{i=1}^{\tilde{l}_1} p_i \rho_i - \sum_{j=1}^{\tilde{l}_2} q_j \sigma_j \right\| - \delta_1 - \delta_2. \end{aligned} \quad (\text{B20})$$

We replaced $\|\sum_{i=\tilde{l}_1+1}^{l_1} p_i \rho_i - \sum_{j=\tilde{l}_2+1}^{l_2} q_j \sigma_j\| \leq \delta_1 - \delta_2$. Let $x_1 = 1 - \delta_1$, $x_2 = 1 - \delta_2$. Then

$$\begin{aligned} &\left\| x_1 \sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - x_2 \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j \right\| = \\ &= \left\| x_1 \sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - x_2 \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j + (x_2 - x_1) \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j \right\| \\ &\leq \left\| \sum_{i=1}^{\tilde{l}_1} p_i \rho_i - \sum_{j=1}^{\tilde{l}_2} q_j \sigma_j \right\| + |\delta_1 - \delta_2| \end{aligned} \quad (\text{B21})$$

Using (B20) we finally get

$$\left\| \sum_{i=1}^{\tilde{l}_1} \tilde{p}_i \rho_i - \sum_{j=1}^{\tilde{l}_2} \tilde{q}_j \sigma_j \right\| \leq \frac{\delta_1 + \delta_2 - |\delta_1 - \delta_2|}{1 - \delta_1}. \quad (\text{B22})$$

Using $a + b + |a - b| = \max\{a, b\}$ and noticing that the same estimate holds, if we exchange δ_1 with δ_2 , we obtain the required estimate.

Appendix C: FOD & CF for $l = 2$

Here we calculate bounds of *FOD* and *CF* in the case when Bob has binary outcomes. By measuring y or y' Bob can create $\{p_i \rho_i\}_{i=0}^1$ or $\{q_j \sigma_j\}_{j=0}^1$ ensembles at Alice's site respectively. Using lemma (1) we know that for all pairs of ρ_i and σ_j

$$\begin{aligned} &\exists \quad \epsilon_{ij} \geq 0, X_{r_0}, X_{r_1}, X_{r_2}, X_{r_3} \quad s.t. \\ &tr(X_{r_0} \rho_0) \geq \epsilon_{00}, \quad \text{and} \quad tr(X_{r_0} \sigma_0) \geq \epsilon_{00} \\ &tr(X_{r_1} \rho_0) \geq \epsilon_{01}, \quad \text{and} \quad tr(X_{r_1} \sigma_1) \geq \epsilon_{01} \\ &tr(X_{r_2} \rho_1) \geq \epsilon_{10}, \quad \text{and} \quad tr(X_{r_2} \sigma_0) \geq \epsilon_{10} \\ &tr(X_{r_3} \rho_1) \geq \epsilon_{11}, \quad \text{and} \quad tr(X_{r_3} \sigma_1) \geq \epsilon_{11} \end{aligned}$$

And FOD and Classical fraction (CF) for a fixed ensemble are given by,

$$\begin{aligned} FOD &= \frac{1}{2k} \max\{ \min\{p_0 \epsilon_{00}, q_0 \epsilon_{00}\}, \min\{p_1 \epsilon_{11}, q_1 \epsilon_{11}\}, \\ &\quad \min\{p_0 \epsilon_{01}, q_1 \epsilon_{01}\}, \min\{p_1 \epsilon_{10}, q_0 \epsilon_{10}\} \} \end{aligned}$$

and

$$\begin{aligned} CF &= \frac{1}{2k} \max\{ \min\{p_0 \epsilon_{00}, q_0 \epsilon_{00}\} + \min\{p_1 \epsilon_{11}, q_1 \epsilon_{11}\}, \\ &\quad \min\{p_0 \epsilon_{01}, q_1 \epsilon_{01}\} + \min\{p_1 \epsilon_{10}, q_0 \epsilon_{10}\} \} \end{aligned}$$

W.l.g. we assume $p_0 \leq q_0 \leq q_1 \leq p_1$. Hence we can rewrite above expressions as

$$FOD = \frac{1}{2k} \max\{p_0 \epsilon_{00}, q_1 \epsilon_{11}, p_0 \epsilon_{01}, q_0 \epsilon_{10}\} \quad (\text{C1})$$

and

$$CF = \frac{1}{2k} \max\{p_0 \epsilon_{00} + q_1 \epsilon_{11}, p_0 \epsilon_{01} + q_0 \epsilon_{10}\} \quad (\text{C2})$$

To simplify calculations one can express ϵ_{ij} in terms of p_i and q_j using lemma(3), i.e., $\epsilon_{ij} \geq 2(1 - \frac{\max\{p_i, q_j\}}{1 - \min\{p_i, q_j\}})$. Then all needed to be done is optimise over p's and q's to calculate FOD and CF . And hence,

$$\begin{aligned}\epsilon_{00} &\geq 2(1 - \frac{\max\{p_1, q_1\}}{1 - \min\{p_1, q_1\}}) \geq 0 \geq 2(1 - \frac{p_1}{q_0}) \\ \epsilon_{01} &\geq 2(1 - \frac{\max\{p_1, q_0\}}{1 - \min\{p_1, q_0\}}) \geq 0 \geq 2(1 - \frac{p_1}{q_1}) \\ \epsilon_{10} &\geq 2(1 - \frac{\max\{p_0, q_1\}}{1 - \min\{p_0, q_1\}}) \geq 2(1 - \frac{q_1}{p_1}) \\ \epsilon_{11} &\geq 2(1 - \frac{\max\{p_0, q_0\}}{1 - \min\{p_0, q_0\}}) \geq 2(1 - \frac{q_0}{p_1})\end{aligned}\quad (C3)$$

Using theorem 2 one can show that $\epsilon_{th} := \max_{ij} \epsilon_{ij} \geq 1/4$. Considering different cases of ϵ_{ij} being equal to ϵ_{th}

$$\begin{aligned}\epsilon_{00} &\geq \frac{1}{4} \rightarrow \max\{\frac{p_0}{4}, 2q_1(1 - \frac{q_0}{p_1}), 0, 2q_0(1 - \frac{q_1}{p_1})\} \\ \epsilon_{01} &\geq \frac{1}{4} \rightarrow \max\{0, 2q_1(1 - \frac{q_0}{p_1}), \frac{p_0}{4}, 2q_0(1 - \frac{q_1}{p_1})\} \\ \epsilon_{10} &\geq \frac{1}{4} \rightarrow \max\{0, 2q_1(1 - \frac{q_0}{p_1}), 0, \frac{q_0}{4}\} \\ \epsilon_{11} &\geq \frac{1}{4} \rightarrow \max\{0, \frac{q_1}{4}, 0, 2q_0(1 - \frac{q_1}{p_1})\}\end{aligned}\quad (C4)$$

Minimizing over all ensemble with the constrains, i.e., $p_0 \leq q_0 \in [0, 1/2]$

$$\begin{aligned}\epsilon_{00} &\geq \frac{1}{4} \rightarrow \\ \min_{p_0 \leq q_0 \in [0, \frac{1}{2}]} [\max\{\frac{p_0}{4}, 2q_1(1 - \frac{q_0}{p_1})\}] &\geq \frac{5 - \sqrt{17}}{8} = 0.1096 \\ \epsilon_{01} &\geq \frac{1}{4} \rightarrow \\ \min_{p_0 \leq q_0 \in [0, \frac{1}{2}]} [\max\{2q_1(1 - \frac{q_0}{p_1}), \frac{p_0}{4}\}] &\geq \frac{5 - \sqrt{17}}{8} = 0.1096 \\ \epsilon_{10} &\geq \frac{1}{4} \rightarrow \\ \min_{p_0 \leq q_0 \in [0, \frac{1}{2}]} [\max\{2q_1(1 - \frac{q_0}{p_1}), \frac{q_0}{4}\}] &\geq 0.1123 \\ \epsilon_{11} &\geq \frac{1}{4} \rightarrow \\ \min_{p_0 \leq q_0 \in [0, \frac{1}{2}]} [\max\{\frac{q_1}{4}, 2q_0(1 - \frac{q_1}{p_1})\}] &\geq 0.1123\end{aligned}$$

Hence taking minimum value of the above four values gives us

$$FOD \geq \frac{1}{2k}(0.1096) \quad (C5)$$

Similarly one can calculate CF which is for this case,

$$CF \geq \frac{1}{2k}(0.1123) \quad (C6)$$